

⑩ 日本国特許庁(JP)

⑪ 特許出願公開

⑫ 公開特許公報(A)

昭63-65499

⑬ Int. Cl.⁴

G 10 L 5/06
3/00

識別記号

3 0 1

庁内整理番号

B-8221-5D
Z-8221-5D

⑭ 公開 昭和63年(1988)3月24日

審査請求 未請求 発明の数 1 (全7頁)

⑮ 発明の名称 構文認識方式

⑯ 特 願 昭61-211169

⑰ 出 願 昭61(1986)9月8日

⑱ 発 明 者 藤 原 祥 隆 神奈川県横須賀市武1丁目2356番地 日本電信電話株式会社
社情報通信処理研究所内

⑲ 発 明 者 菊 池 英 夫 神奈川県横須賀市武1丁目2356番地 日本電信電話株式会社
社情報通信処理研究所内

⑳ 発 明 者 中 村 修 神奈川県横須賀市武1丁目2356番地 日本電信電話株式会社
社情報通信処理研究所内

㉑ 出 願 人 日本電信電話株式会社 東京都千代田区内幸町1丁目1番6号

㉒ 代 理 人 弁理士 鈴木 誠

明 細 書

1. 発明の名称

構文認識方式

2. 特許請求の範囲

(1) 音声により入力された自然言語文の構文を、音声パラメータ列から直接認識する方式において、キーワードを特徴づける音声パラメータ列を格納したキーワード辞書部と、入力音声パラメータ列と当該キーワード辞書に格納された音声パラメータ列を照合し、該音声パラメータ列が存在する確からしさを表すスコアを計算するワードスポッティング部と、該ワードスポッティング部により計算された各キーワードのスコアと認識対象とする構文候補について予め指定されたキーワードとに基づき構文候補が存在するか否かを評価する構文スコアリング部とを備え、前記キーワード辞書部は、構文候補を特徴づけるキーワードカテゴリ単位に分割し構成するとともに、各々のキーワードカテゴリにはそのカテゴリの属するキーワードの音声パラメータ列を格納し、前記ワードスポッ

ティング部は、該キーワードカテゴリを単位にワードスポッティング処理を行うことを特徴とする構文認識方式。

(2) 前記構文スコアリング部は、キーワードカテゴリの種別、出現順序ならびに出現範囲に関する構文知識に基づき、構文候補を特徴づけるキーワードカテゴリのスコアを加重加算することにより構文候補の存在の確からしさを表す構文スコアを求めることを特徴とする特許請求の範囲第1項記載の構文認識方式。

(3) 前記構文スコアリング部は、構文候補を特徴づけるキーワードカテゴリのスコアに基づき求めた構文候補の存在の確からしさを表す構文スコアを、該構文スコアの到達可能な上限値により正規化した相対値を、該構文の構文スコアとすることを特徴とする特許請求の範囲第1項記載の構文認識方式。

(4) 前記ワードスポッティング部は、構文候補を特徴づける前記キーワードカテゴリ単位に少なくとも1個を備え、着目する入力音声パラメータ列

に対して、並列にワードスポッティング処理を行うことを特徴とする特許請求の範囲第1項記載の構文認識方式。

3. 発明の詳細な説明

〔産業上の利用分野〕

本発明は、構文認識方式に係り、詳しくは音声により入力された自然言語文（例えば日本語文）の構文を、音声パラメータ系列（特徴パラメータ系列）から直接認識する方式に関する。

〔従来の技術〕

従来、音声パラメータ列の段階で、入力された構文がなにかを認識する手法としてキーワードスポッティング手法が知られている。本手法は、構文候補を特徴づける単語（キーワード）の音声パラメータ列（例えば、一定時間間隔毎に入力音声信号をサンプリングして得た相関係数の列）をあらかじめ辞書に登録しておき、入力された音声パラメータ列とこれらのキーワードの音声パラメータ列を比較して評価点づけ（スコアリングという）を行うことにより、入力音声パラメータ列中に着

目するキーワードが存在するか否かを判定し、入力された構文を認識するというものである。これについては、例えば、W.A.Lee氏が“The Value of Speech Recognition Systems”と題してTRENDS IN SPEECH RECOGNITION, 1980, PRENTICE-HALL, INC. pp.10-11で論じている。

以下に日本語による切符予約質問応答システムを想定して、キーワードスポッティング（以後、単にワードスポッティングと記す）を用いた従来の構文認識方式について説明する。

第4図は従来の構文認識方式を適用した切符予約質問応答システムを示したもので、ここでは説明の簡単化のため構文候補は“切符予約希望”と“値段照合”の二つとしている。システムはキーワード辞書1とワードスポッティング部2と構文判定部3からなる。キーワード辞書1には各構文候補を特徴づけるキーワードに対応する音声パラメータ列を予め登録しておく。ワードスポッティング部2は入力音声パラメータ列にふくまれている

キーワードの抽出を行う。構文判定部3はワードスポッティング部2で抽出したキーワードにもとづいて構文候補を決定する。入力音声パラメータ列は、例えば入力音声を一定周波でサンプリングして得た相関係数、LPCケプストラム等の列で示される。以下ワードスポッティング部2を中心に第4図の動作を説明する。

- ① キーワード辞書1から、あるキーワードの音声パラメータ列を取り出しキーワードレジスタ23に設定する。
- ② 対象とする入力音声パラメータ列の先頭からキーワードレジスタ23に設定された音声パラメータ列の長さに等しい長さの音声パラメータ部分列を入力レジスタ22に設定し、キーワードスコアリング部24で両者の間の照合を行い、類似度に対応したスコア（通常、距離計算と呼ぶ計算により求めた値）を求め、その値を保持する。距離計算は、例えば
$$\sum_{j=1}^n (R_j - S_j)^2$$
 として示される。R_jは入力レジスタ22の位置j

にあるパラメータ値、S_jはキーワードレジスタ23の位置jにあるパラメータ値である。

- ③ 上記②の操作を入力音声パラメータ列の部分列の開始位置を一つずつずらしながら繰り返し実行し、入力音声パラメータ列の全ての部分列と着目しているキーワードの音声パラメータ列の間の類似度に関するスコアを求め、その中の最良のスコアがある基準を満足するときに着目するキーワードが存在すると判定する。
- ④ 上記①～③の操作をキーワード辞書1に登録されたすべてのキーワードについて行い、入力音声パラメータ列にふくまれているとみなされるキーワード候補を検出する。
- ⑤ 構文判定部3では構文毎にその構文を特徴づけるものとして予め与えられたキーワードセットの情報とワードスポッティング部2により検出されたキーワードのセットを比較し、予め指定された基準（例えば、指定されたキーワードの50%が存在する等）を満足する構文候補を選択し出力する。

(発明が解決しようとする問題点)

上記従来の構文認識方式においては、着目する構文候補の数が増えると辞書に登録するキーワード数が増し、これに伴い類似の特徴パラメータをもつキーワードが増す。このため、キーワードスコアリング部では、対象とする入力文中に着目する単語が存在しないのに存在すると検出したり、或いはある単語を他の単語と取り違えて認識するなどの誤りが増加するとともに、これらの誤りは構文間にまたがって起こるため、構文を正しく認識する確率（こゝではこれを構文認識性能と呼ぶ）が低下する欠点がある。

また、上述した構文認識方式は、①基本的にはキーワードの有無を問題とするデジタル的検出方法をとるため、着目している構文候補がどの程度入力文に似ているかという類似の程度を捉えない、②キーワードの出現位置、或いはキーワードの出現順序等の構文を特徴づける情報に着目していないことにより、キーワードは同じであるがその順序が異なる構文候補間の区別がつかない

構成される。ワードスポッティング部は、当該キーワードカテゴリ単位に、そのカテゴリに登録されたキーワードと入力音声パラメータ列との間でワードスポッティングを行う。その結果、ワードスポッティングで使用する辞書の規模は構文候補の数とは無関係にほぼ一定となるため、これにより一定水準の構文認識性能が維持される。さらにキーワード辞書は上記のような排他性を有するため、ワードスポッティングにおける誤りの影響は、ある特定のキーワード辞書内すなわち構文候補を特徴付ける特定キーワードカテゴリの範囲内に限定され、その結果、誤りは高々その構文のスコアの多少に反映されるだけに局所化される。

また、構文スコアリング部は、各構文候補に関し、どのキーワードがあるべきか、各キーワードはどの範囲にあるべきか、各キーワードの出現順序はどうあるべきか、といった構文知識が予め与えられており、これらの知識に基づいてワードスポッティング部で計算されたキーワードのスコア値（アナログ値）を加加重算する。さらに構文候

(例 ・切符を予約したい、 ・予約した切符を取消したい)等のため、類似な表現を含む多数の構文候補の中から正解構文を精度良く固定することは困難と考えられる。実際、本技術を構文の予備的選択手段として位置付けている文献がある。

(例えば、新見著、“音声認識”，共立出版，p 184)

本発明の目的は、構文候補の多少によらず一定水準の構文認識性能を維持でき、しかも類似な表現を含む多数の構文候補の中から正解構文候補を精度良く検出することを可能とする新しい構文認識方式を提供することにある。

(問題点を解決するための手段及び作用)

本発明の構文認識方式の主要構成要素は、キーワード辞書と、当該辞書を用いるワードスポッティング部と、入力文と構文候補の類似性を数値的に評価するスコアリング部とからなる。

キーワード辞書は構文候補単位に排他的に、かつ同じ構文候補内においても該構文候補を特徴づけるキーワードカテゴリ単位に排他的に分割して

補の違いによるキーワード組数の違いを考慮し、得られたスコア値と当該構文について到達可能なスコアの上限値との相対比（正規化されたスコア値）をとり、これを着目する構文候補のスコア値とする。このような特徴を持った構文スコアリング法を導入することにより、類似な表現を含む多数の構文候補の中から正解を精度良く検出することが可能となる。

(実施例)

以下、本発明の一実施例について図面により説明する。なお、以下の実施例でも日本語による切符予約質問応答システムを想定して、構文候補は“切符予約希望”と“値段照会”の二つとし、第4図と同じ例を使用することにする。

第1図は本発明の一実施例のシステム構成を示す。本構文認識システムは、キーワードが登録されたキーワード辞書部1、ワードスポッティングを行うワードスポッティング部2、各構文候補のスコアを計算し、最良スコアの構文候補を選択する構文スコアリング・判定部3より構成される。

キーワード辞書1は第2図に示すように、構文候補毎に、また構文候補内ではその構文を特徴付けるキーワードカテゴリ毎に相互に排他的に構成される。即ち、第1図の実施例では、切符予約を希望する構文候補に関しては、予約希望の意思表示を表すキーワードを集めた“予約希望”カテゴリと、切符を意味するキーワードをあつめた“切符”カテゴリの2種類のキーワード辞書11, 12が構成される。これは、例えば「チケットを予約したい」という構文の入力を想定している。同様にして、値段を照会する構文候補に関しては、値段照会の意思表示を表すキーワードをあつめた“値段”カテゴリと、値段を表すキーワードをあつめた“値段”カテゴリの2種類のキーワード辞書13, 14が構成される。これは、例えば「A席の値段を知りたい」という構文の検出を想定したものである。

ワードスポッティング部2は、ワードスポッティングの単位であるキーワードカテゴリを選択する選択回路21、入力音声パラメータ列の部分列

を設定する入力レジスタ22、辞書1に登録されているキーワードの音声パラメータ列を設定するキーワードレジスタ23、キーワードレジスタ23に設定されたキーワードに対し、最も類似した部分列を入力音声パラメータ列からみつけ、そのスコアを保持するキーワードスコアリグ部24よりなる。このワードスポッティング部2では、選択回路21が順次キーワードカテゴリを一つずつ選択し、キーワードスコアリグ部24が当該カテゴリに登録されている全てのキーワードに関しワードスポッティングを行い、それらのスコアを保持する。なお、ワードスポッティングの方法は第4図の従来技術で説明したと同じ方法を使用する。

構文スコアリング・判定部3は、上述のように各構文毎に、どのキーワードがあるべきか、各キーワードはどの範囲に存在すべきか、各キーワードの出現順序はどうあるべきか、といった構文を特徴付ける情報（構文知識）が予め与えられており、その動作は以下ようになる。まず最初の構文候補（この場合は、切符予約希望）に注目し、

指定されたキーワードカテゴリ（予約希望、切符）に関し計算されたスコア値をキーワードスコアリグ部24より読み取る。つぎに、与えられた構文知識に基づき、これらのスコア値を加重加算する。例えば、 $(\sum W1j \cdot W2j \cdot Aj) \cdot W3$ のような加重計算を行う。ここで、 Aj はキーワードカテゴリjのスコア、 $W1j$ はキーワードカテゴリjが構文決定要因としてもつ重み（例えば、 $W11=1$, $W12=0.5$ ）、 $W2j$ はキーワードカテゴリjが検出された位置に関する重み（例えば、指定範囲で検出されたときは $W2j=1$ 、指定範囲外のときは $W2j=0.3$ ）、 $W3$ はキーワードの順序に関する重み（例えば、検出されたキーワードが指定順序にあるとき $W3=1$ 、指定順序にないとき $W3=0.2$ ）である。さらに得られた加重加算結果と当該構文候補に関する加重加算結果の到達可能な上限値との相対比をとり、これを当該構文候補の構文スコアとする。これらの操作をすべての構文候補について実行し、各構文候補の構文スコアを求める。最後にこれらの構文ス

コアの中で最良のスコアを持つ構文候補を選び、正解構文候補として出力する。

第3図に本発明の他の実施例のシステム構成を示す。上記第1図の構成の場合、計算時間はワードスポッティング処理で支配される。第3図は、このワードスポッティング処理の高速化を図ったものである。即ち、第3図は第1図のワードスポッティング処理をキーワードカテゴリ単位に並列化したものであり、これにともない第1図の選択回路21が除去されている。また、構文スコアリングを構文候補について並列に行うことを目的に、第1図の構文スコアリング・判定部3は構文スコアリング部61, 62と構文判定部7に機能を分割している。

第3図において、ワードスポッティングおよび構文スコアリングは第1図と同様に行うが、入力音声パラメータ列に対して、ワードスポッティング51～54では割当られたキーワードカテゴリに関してそれぞれ並列にワードスポッティング処理を行い、得られたキーワードのスコアをそれぞ

れキーワードスコアリング部に保持する。次に構文スコアリング部61、62はそれぞれ並列に、指定されたキーワードカテゴリに関するスコア値を対応するキーワードスコアリング部より読みだし、加重加算、正規化処理を行い構文スコアを求める。最後に構文判定部7は各構文候補について得られた構文スコアの中の最良スコアを示す構文候補を選択し、これを正解として出力する。

なお、第3図の考えを発展させることにより、ワードスポッティング部は最大、各ワードスポッティング部がキーワード1個を分担する個数まで並列化可能であり、ワードスポッティング処理をさらに高速化することが可能になる。

〔発明の効果〕

以上説明したように、本発明によれば、排他的に分割し構成したキーワード辞書を用することにより、ワードスポッティングで使用する辞書の規模は構文候補の数とは無関係にほぼ一定となり、これにより一定水準の構文認識性能が維持される。さらにキーワード辞書の排他性により、ワードス

ポッティングにおける誤りの影響が特定のキーワードカテゴリの範囲内に限定されるため、誤りが高々その構文のスコアの多少に反映されるだけに局所化され、誤りに対しフェールセーフなシステムとなる。

また、構文知識に基づくキーワードのスコア値の加重加算と、構文候補毎のスコア上限値による加重加算結果の正規化により、類似な表現を含む多数の構文候補から正解構文を精度良く検出することが可能となる。さらに、このように精度の良い構文認識が可能となると、構文解析、意味解析等の自然言語処理技術の利用を著しく簡略化でき或いは不要とすることができ、情報処理装置の音声を用いるマンマシンインターフェイスを経済的にかつ能率良く（構文認識の高速化）構成できる。

4. 図面の簡単な説明

第1図は本発明の構文認識方式の一実施例を示す図、第2図は本発明によるキーワード辞書構成を説明するための図、第3図は本発明の構文認識

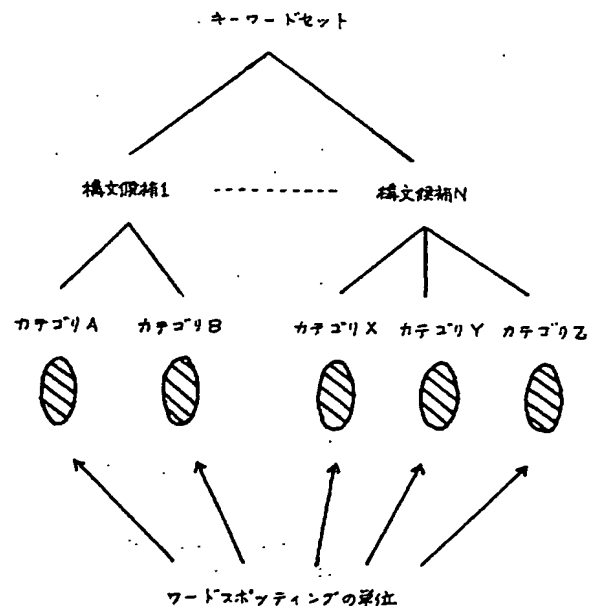
方式の他の実施例を示す図、第4図はワードスポッティングを用いた従来の構文認識方式を説明するための図である。

- 1…キーワード辞書、 2…ワードスポッティング部、 3…構文スコアリング・判定部、
51～54…ワードスポッティング部、
61、62…構文スコアリング部、
7…構文判定部。

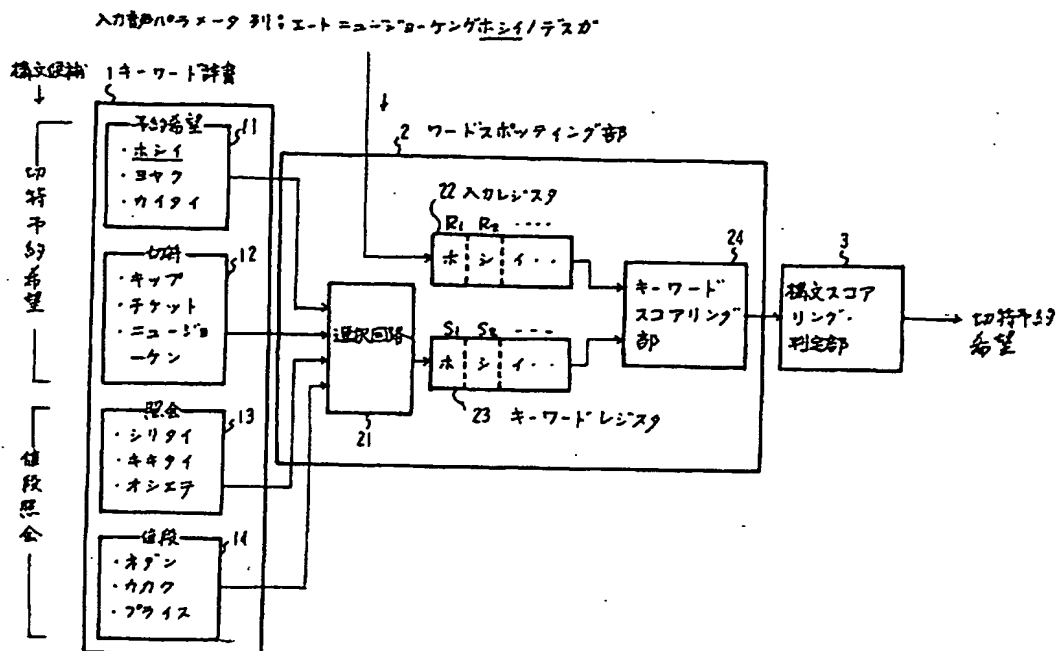
代理人弁理士 鈴木



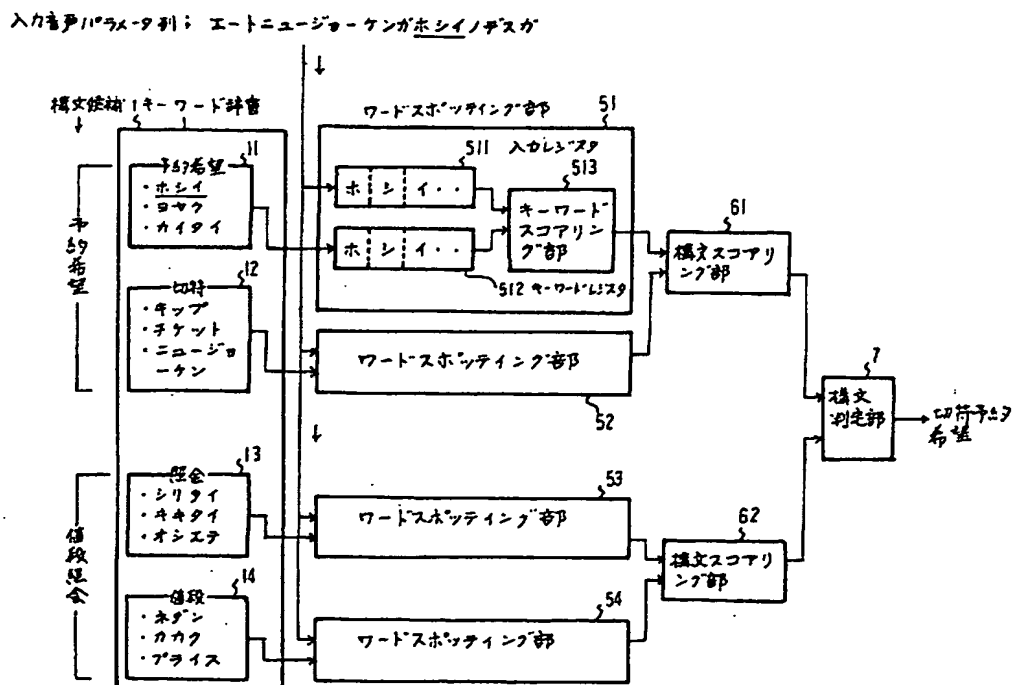
第2図



第 1 図



第 3 图



第4図

入力音声パラメータ列: エトニューシヨウゲンガホシノテスガ

